Journal of Integer Sequences, Vol. 20 (2017),
Article 17.3.7

A Combinatorial Identity Concerning Plane Colored Trees and its Applications

Ricky X. F. Chen and Christian M. Reidys
Biocomplexity Institute and Department of Mathematics
Virginia Polytechnic Institute and State University
1015 Life Science Circle
Blacksburg, VA 24061
USA
cxiaof6@vt.edu
duck@santafe.edu

Abstract

In this note, we obtain a combinatorial identity by counting some colored plane trees. This identity has a plethora of implications. In particular, it solves a bijective problem in Stanley's collection "Bijective Proof Problems", and gives a formula for the Narayana polynomials, as well as an equivalent expression for the Harer-Zagier formula enumerating unicellular maps.

1 Introduction

This note is motivated by giving a combinatorial proof for the following bijective problem in Stanley's collection "Bijective Proof Problems" [8, (15)]:

$$\sum_{k=0}^{n} \binom{n}{k}^2 x^k = \sum_{j=0}^{n} \binom{n}{j} \binom{2n-j}{n} (x-1)^j. \tag{1}$$

Identities involving the Narayana numbers $N_{n,k} = \frac{1}{n}\binom{n}{k}\binom{n}{k-1}$ [11, A001263] have been well studied; for instance, see the work [2, 6, 7]. The authors found that the new expression of the Narayana polynomials obtained by Mansour and Sun [7] (and independently by Chen and Pang [2]) is related to (1). The new expression of the Narayana polynomials reads

$$\sum_{k=1}^{n} N_{n,k} y^k = \sum_{k=0}^{n} \frac{1}{n+1}\binom{n+1}{k}\binom{2n-k}{n}(y-1)^k. \tag{2}$$

This note is organized as follows. In Section 2, we obtain an elementary identity by counting some kind of colored plane trees. This identity implies the identities (1) and (2) as special cases. Furthermore, it gives a new expression for the well-known Harer-Zagier formula, i.e., the generating polynomials for unicellular maps [5]. The new expression allows us to present a new explicit formula for the numbers $A(n,g)$ [11, A035309] counting unicellular maps having n edges and genus g [3] and achieve the most recent advance on the subject, i.e., Chapuy's recursion for $A(n,g)$, in a different way [3]. In Section 3, we enumerate some variations of the colored plane trees so that additional binomial identities are obtained.

2 A fundamental identity and consequences

In this section, we prove the following identity:

Theorem 1. *For $n, q \geq 0$, $c \in \mathbb{C}$, we have*

$$\sum_{k=0}^{n} \binom{n}{k}\binom{2n+c+q-k-2}{n+q-1} z^k = \sum_{k=0}^{n} \binom{n}{k}\binom{n+c+q-2}{k+q-1}(z+1)^k. \tag{3}$$

A *colored-labeled plane tree* with $n+1$ vertices is a plane tree where its vertices are uniquely labeled by $[n+1] = \{1, 2, \ldots, n+1\}$ and its leaves are either not colored, or colored N or Y. A vertex which is not a leaf is an *internal* vertex. Let $\mathrm{int}(T)$, $\mathrm{lev}_Y(T)$ and $\mathrm{lev}_N(T)$ denote the set of the internal vertices, Y-leaves and N-leaves in a colored-labeled plane tree T, respectively. As usual, the cardinality of a set S is denoted by $|S|$.

To begin, we recall a bijection between labeled plane trees and sets of matches, called *Chen's bijective algorithm* [1]. A *match* is a labeled plane tree with two vertices, i.e., consists of a root and a leaf.

Proposition 2. *(Chen [1]) There is a bijection ϕ from labeled plane trees with labels in $[n+1]$ to sets of n matches with labels in $\{1, \ldots, n+1, (n+2)^*, \ldots, (2n)^*\}$. For a labeled plane tree T, every internal vertex in T will appear as the root of some match in $\phi(T)$, while every leaf of T will appear as the leaf of some match in $\phi(T)$, and vice versa.*

Note there are n matches in $\phi(T)$ in total. Except these matches having (unstarred) roots which correspond to internal vertices, every match of the rest must have a starred root (i.e., with a label marked with *). For details with respect to *Chen's bijective algorithm*, we refer the reader to Chen [1].

Let $\Gamma_{n,c,q}$ denote the set of colored-labeled plane trees on $[n+c+q]$ in which all vertices with labels in $[q]$ are all uncolored leaves, and all vertices with labels in $\{q+1, \ldots, q+c\}$ are internal. Using Proposition 2, we obtain

Lemma 3. *There is a bijection between the set of colored-labeled plane trees $T \in \Gamma_{n,c,q}$ with $|\,int(T)| + |\,lev_Y(T)| = k + c$ and the set of pairs (A, χ) where $A \subseteq [n + c + q] \setminus [c + q]$ with $|A| = k$ and χ is a set of $n + c + q - 1$ matches with labels in $\{1, \ldots, n + c + q, (n + c + q + 1)^*, \ldots, 2(n+c+q-1)^*\}$ such that all vertices with labels in $\{q+1, \ldots, q+c\}$ are roots and other unstarred roots of χ are in A.*

Proof. Given $T \in \Gamma_{n,c,q}$ with $|\,\mathrm{int}(T)| + |\,\mathrm{lev}_Y(T)| = k + c$, clearly the summation of the number of internal vertices other than those in $\{q + 1, \ldots, q + c\}$ and the number of Y-leaves is k. Let A denote the union of this set of internal vertices and Y-leaves. According to Proposition 2, without considering the coloring of leaves, T corresponds to a set of $n+c+q-1$ matches with labels in $\{1, \ldots, n + c + q, (n + c + q + 1)^*, \ldots, 2(n + c + q - 1)^*\}$ such that all vertices with labels in $\{q + 1, \ldots, q + c\}$ are roots and any other unstarred root of χ is contained in A. Accordingly, T corresponds to the pair (A, χ).

Conversely, given a pair (A, χ), the set of matches χ corresponds to a tree $T \in \Gamma_{n,c,q}$ according to Proposition 2. It thus remains to color the leaves of T. Note, there are three classes of leaves: those in $[q]$ which are uncolored by assumption, those in A and those not in A. We color those leaves in A color Y and those not in A color N. Thus, vertices in A are either internal or Y-leaves. Hence, $|\,\mathrm{int}(T)| + |\,\mathrm{lev}_Y(T)| = k + c$, completing the proof. □

Figure 1 illustrates the bijection.

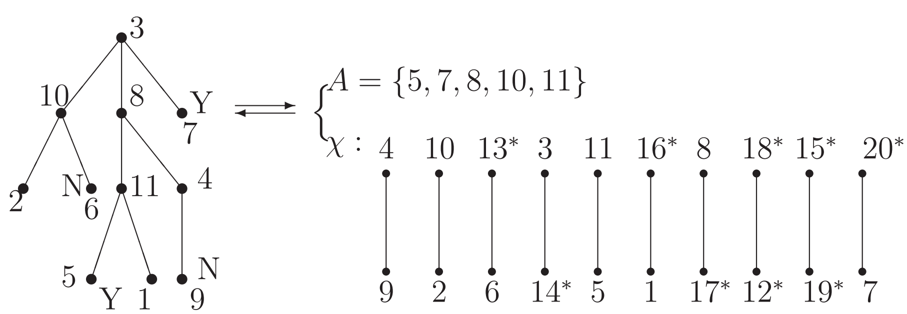

Figure 1: A tree in $\Gamma_{11,2,2}$ and its corresponding pair (A, χ).

Proof of Theorem 1. firstly, we claim that the number of trees $T \in \Gamma_{n,c,q}$ such that $|\,\mathrm{int}(T)|+ |\,\mathrm{lev}_Y(T)| = k + c$ is

$$\binom{n}{k}\binom{k+n+c+q-2}{n+q-1}(n+c+q-1)!.$$

Based on Lemma 3, there are $\binom{n}{k}$ ways to choose A. Besides those c prescribed roots in $\{q+1, \ldots, q+c\}$, the rest of $(n+c+q-1)-c = n+q-1$ roots can only come from the set $A \cup \{(n+c+q+1)^*, \ldots, 2(n+c+q-1)^*\}$ so that there are $\binom{k+n+c+q-2}{n+q-1}$ ways to determine the rest of $n+q-1$ roots of χ. At last, there are $(n+c+q-1)!$ ways of pairing up all roots and leaves to obtain $n + c + q - 1$ matches, whence the claim.

Next we weigh each tree $T \in \Gamma_{n,c,q}$ by $z^{|\operatorname{lev}_N(T)|}$. Then, the total weight over all trees in $\Gamma_{n,c,q}$ is

$$\sum_{k=0}^{n} z^{n-k} \binom{n}{k} \binom{k+n+c+q-2}{n+q-1} (n+c+q-1)!.$$

Counting in a different way, we inspect that the total weight over all trees $T \in \Gamma_{n,c,q}$ with $|\operatorname{int}(T)| = c+k$ equals

$$\binom{n}{k} \binom{n+c+q-2}{n+q-1-k} (n+c+q-1)!(z+1)^{n-k}.$$

It follows from Proposition 2 that a tree $T \in \Gamma_{n,c,q}$ with $|\operatorname{int}(T)| = c+k$ (without considering the coloring of leaves) corresponds to a set of $n+c+q-1$ matches χ where there are $c+k$ unstarred roots in total. Since vertices in $\{q+1, \dots, q+c\}$ are always unstarred roots of χ by assumption, there are $\binom{n}{k}$ ways to choose from $\{q+c+1, \dots, n+q+c\}$ k additional unstarred roots of χ. Furthermore, there are $\binom{n+c+q-2}{(n+c+q-1)-c-k}$ ways to choose starred roots and $(n+c+q-1)!$ different ways of pairing up. Finally, all $n-k$ leaves other than those in $[q]$ can be either colored Y or N, whence each of them contributes a weight of $(z+1)$. Summing over all $0 \le k \le n$, we also obtain the total weight over all trees in $\Gamma_{n,c,q}$

$$\sum_{k=0}^{n} \binom{n}{k} \binom{n+c+q-2}{n+q-1-k} (n+c+q-1)!(z+1)^{n-k}.$$

Accordingly we arrive at the identity

$$\sum_{k=0}^{n} z^{n-k} \binom{n}{k} \binom{k+n+c+q-2}{n+q-1} = \sum_{k=0}^{n} \binom{n}{k} \binom{n+c+q-2}{n+q-1-k} (z+1)^{n-k}.$$

Since both sides of (3) can be viewed as polynomials in c, it holds for any $c \in \mathbb{C}$, completing the proof of Theorem 1. $\square$

Theorem 1 can be proved alternatively, e.g., directly working with matchings instead of plane trees. However, to the best of our knowledge, this was not presented elsewhere.

As the first application, it can solve the bijective proof problem of Stanley [8, (15)]:

Corollary 4 (Stanley [8, (15)])**.** *For $n \ge 0$, we have*

$$\sum_{k=0}^{n} \binom{n}{k}^2 z^k = \sum_{j=0}^{n} \binom{n}{j} \binom{2n-j}{n} (z-1)^j. \tag{4}$$

Proof. Applying the same combinatorial argument of Theorem 1 to the set $\Gamma_{n,1,1}$ (with z replaced by $z-1$) completes the proof. $\square$

Second we obtain a combinatorial proof for the new expression of the Narayana polynomials obtained in Chen and Pang [2] as well as Mansour and Sun [7] via studying statistics of lattice paths:

Corollary 5. *For $n \geq 0$, the Narayana numbers $N_{n,k} = \frac{1}{n}\binom{n}{k}\binom{n}{k-1}$ satisfy*

$$\sum_{k=1}^{n} N_{n,k} z^k = \sum_{k=0}^{n} \frac{1}{n+1}\binom{n+1}{k}\binom{2n-k}{n}(z-1)^k. \tag{5}$$

Proof. Applying the same combinatorial argument of Theorem 1 to the set $\Gamma_{n,2,0}$ (with z replaced by $z-1$) completes the proof. □

It is well-known that, the large Schröder numbers (A006318) S_n [4, 9] which counts the number of plane trees having n edges with leaves colored by one of two colors (say color Y and color N), equals the evaluation at $z = 2$ in the n-th Narayana polynomial, i.e.,

$$S_n = \sum_{k=1}^{n} N_{n,k} 2^k.$$

In particular, for $q = 0, x = 2$, i.e., $\Gamma_{n,2,0}$, Theorem 1 implies the following theorem.

Theorem 6. *Let T_{n+1} denote the number of plane trees of $n+1$ edges with 2 different internal, marked vertices and bi-colored leaves. Then, $T_{n+1} = \binom{n+1}{2} S_n$.*

Proof. From the proof of Theorem 1, we know that the number of trees in $\Gamma_{n,2,0}$ is given by

$$\sum_{k=0}^{n} \binom{n}{k}\binom{n}{k+1}(n+1)!2^k.$$

Ignoring labels we obtain the number of (unlabelled) plane trees of $n+1$ edges with 2 different internal, marked vertices and bi-colored leaves to be

$$\frac{1}{2!n!}\sum_{k=0}^{n} \binom{n}{k}\binom{n}{k+1}(n+1)!2^k = \frac{(n+1)n}{2}\sum_{k=1}^{n} N_{n,k}2^k = \binom{n+1}{2} S_n,$$

which completes the proof. □

Since the large (and small) Schröder numbers satisfy the recurrence [4]

$$3(2n-1)S_{n-1} = (n+1)S_n + (n-2)S_{n-2}, \quad n \geq 2 \tag{6}$$

and $S_0 = 1,\ S_1 = 2$, we have the following corollary.

Corollary 7. *The numbers $(T_n)_{n\geq 1}$ satisfy*

$$3(2n-1)T_n = (n-1)T_{n+1} + nT_{n-1}. \tag{7}$$

Based on Eq. (3), we can also give a new expression for the generating polynomials of unicellular maps. A *unicellular map* is a graph embedded in a closed orientable surface such that the complement is homeomorphic to an open disk. The number of handles of the surface is called the *genus* of the unicellular map. Let $A(n, g)$ denote the number of unicellular maps of genus g with n edges. The Harer-Zagier formula [5] gives a generating polynomial for these numbers, which reads

$$\sum_{g\geq 0} A(n,g)x^{n+1-2g} = \frac{(2n)!}{2^n n!}\sum_{k\geq 1} 2^{k-1}\binom{n}{k-1}\binom{x}{k}. \tag{8}$$

Now we can give a new expression via Theorem 1:

Corollary 8.

$$\frac{(2n)!}{2^n n!}\sum_{k\geq 1}\binom{n}{k-1}\binom{x}{k}2^{k-1} = \frac{(2n)!}{2^n n!}\sum_{k\geq 0}\binom{n}{k}\binom{x+n-k}{n+1}. \tag{9}$$

Proof. Setting $q = 2, x = x - n, z = 1$ in (3) completes the proof. □

From the RHS of Eq. (9), we can obtain a new explicit formula for $A(n, g)$ in terms of a convolution of the Stirling numbers of the first kind $C(n, k)$ (A132393) and a new way to obtain Chapuy's recursion [3].

3 A variation

In this section, we count a special subclass of colored plane trees and obtain further identities.

Theorem 9. *For $0 \leq n, 1 \leq k, 0 \leq t < k, q \in \mathbb{Z}, x \in \mathbb{C}, \omega_k = e^{\frac{j2\pi}{k}}, j^2 = -1$, we have*

$$\sum_{l=0}^{n}\binom{kn+t}{kl+t}\binom{kl+x+q+kn+2t-2}{q+kn+t-1}z^{kl} = \frac{1}{k}\sum_{i=0}^{kn+t}\binom{kn+t}{i}\binom{x+q+kn+t-2}{q+kn+t-1-i}z^{i-t}\sum_{l=1}^{k}\frac{(1+z\omega_k^l)^{kn+t-i}}{(\omega_k^l)^{t-i}}. \tag{10}$$

Proof. Considering the trees $T \in \Gamma_{kn,x,q}$ with $|\operatorname{int}(T)| + |\operatorname{lev}_Y(T)| = x + kl + t, l \geq 0$, we obtain, along the lines of Theorem 1

$$\sum_{l=0}^{n}\binom{kn+t}{kl+t}\binom{kl+kn+x+q+2t-2}{kn+t+q-1}z^{kn-kl} = \sum_{i=0}^{kn+t}\binom{kn+t}{i}\binom{kn+t+x+q-2}{kn+t+q-i-1}\sum_{kl+t\geq i}\binom{kn+t-i}{kl+t-i}z^{kn-kl}.$$

Canceling the term z^{kn} and setting $z = z^{-1}$ in the above identity, the second summation on the right hand side becomes

$$z^{i-t} \sum_{l \geq 0} \binom{kn+t-i}{kl+t-i} z^{kl+t-i}.$$

From the identity [10, Eq. (1.53)], we have

$$\sum_{i \geq 0} \binom{n}{a+ki} x^{a+ki} = \frac{1}{k} \sum_{i=1}^{k} (\omega_k^i)^{-a} (1 + x\omega_k^i)^n, k-1 \geq a, a \in \mathbb{Z}.$$

Therefore, setting $a = t-i, n = kn+t-i, k = k$ we obtain

$$\sum_{l \geq 0} \binom{kn+t-i}{kl+t-i} z^{kl+t-i} = \frac{1}{k} \sum_{l=1}^{k} (\omega_k^l)^{i-t} (1 + z\omega_k^l)^{kn+t-i}$$

and the proof follows. $\square$

As the first application of Theorem 9, we obtain two identities involving the Narayana numbers.

Corollary 10.

$$\sum_{l} N_{n,l+1} z^l (1+z)^{n-l} = \sum_{l} \frac{1}{n} \binom{n}{l} \binom{n+l}{l-1} z^l, \tag{11}$$

$$\sum_{l=0}^{2n+s_1} N_{2n+s_1,l+1} 2^{2n+s_1-l} = \sum_{l=0}^{n} \frac{1}{n} \binom{2n+s_1}{2l+s_1} \binom{2n+2l+2s_1}{2l+s_1+1}. \tag{12}$$

Proof. Setting $k = 1$ in Eq. (10), we obtain the first identity. Setting $k = 2,\ q = 0,\ x = 2,\ z = 1$ leads to the second one. $\square$

Furthermore, we derive the following relations which can be viewed as analogues of the well-known relation:

$$\sum_{i=1,\ i \in odd}^{n} \binom{n}{i} = \sum_{i=0,\ i \in even}^{n} \binom{n}{i}.$$

Corollary 11. *For $n \geq 0$, we have*

$$\sum_{l=0}^{n} \binom{2n}{2l} \binom{2l+2n}{2l} = \sum_{l=0}^{n-1} \binom{2n}{2l+1} \binom{2n+2l+1}{2l+1} + 1, \tag{13}$$

$$\sum_{l=0}^{n} \binom{2n+1}{2l} \binom{2l+2n+1}{2l} = \sum_{l=0}^{n} \binom{2n+1}{2l+1} \binom{2n+2l+2}{2l+1} - 1. \tag{14}$$

Proof. Setting $x = q = 1,\ z = 1,\ k = 2$ in Eq. (10), we obtain, for $0 \leq s_1 < 2$,

$$\sum_{l=0}^{n} 2\binom{2n+s_1}{2l+s_1}\binom{2l+2n+2s_1}{2n+s_1} - 1 = \sum_{l=0}^{2n+s_1} \binom{2n+s_1}{l}^2 2^{2n+s_1-l}.$$

However, from Eq. (4), we have

$$\sum_{l=0}^{2n+s_1} \binom{2n+s_1}{l}^2 2^{2n+s_1-l} = \sum_{j=0}^{2n+s_1} \binom{2n+s_1}{j}\binom{2n+s_1+j}{j}.$$

Then,

$$\sum_{l=0}^{n} 2\binom{2n+s_1}{2l+s_1}\binom{2l+2n+2s_1}{2l+s_1} - 1 = \sum_{j=0}^{2n+s_1} \binom{2n+s_1}{j}\binom{2n+s_1+j}{j},$$

from which the corollary follows. □

4 Acknowledgments

The authors thank the anonymous referees for their valuable feedback.

References

[1] W. Y. C. Chen, A general bijective algorithm for trees, *Proc. Natl. Acad. Sci. USA* **87** (1990), 9635–9639.

[2] W. Y. C. Chen and S. X. M. Pang, On the combinatorics of the Pfaff identity, *Discrete Math.* **309** (2009), 2190–2196.

[3] R. X. F. Chen and C. M. Reidys, New formulas counting one-face maps and Chapuy's recursion, in revision.

[4] D. Foata and D. Zeilberger, A classic proof of a recurrence for a very classical sequence, *J. Combin. Theory Ser. A* **80** (1997), 380–384.

[5] J. Harer and D. Zagier, The Euler characteristic of the moduli space of curves, *Invent. Math.* **85** (1986), 457–485.

[6] N. Y. Li and T. Mansour, An identity involving Narayana numbers, *European J. Combin.* **29** (2008), 672–675.

[7] T. Mansour and Y. Sun, Identities involving Narayana polynomials and Catalan numbers, *Discrete Math.* **309** (2009), 4079–4088.

[8] R. P. Stanley, Bijective proof problems, 2009, `http://www-math.mit.edu/~rstan/bij.pdf`.

[9] L. W. Shapiro and R. A. Sulanke, Bijections for the Schröder numbers, *Math. Mag.* **73** (2000), 369–376.

[10] R. Sprugnoli, Riordan array proofs of identities in Gould's book, 2006, `http://www.dsi.dsi.unifi.it/~resp/GouldBK.pdf`.

[11] N. J. A. Sloane, The On-line Encyclopedia of Integer Sequences, `http://oeis.org`.

2010 *Mathematics Subject Classification*: Primary 05A19.
Keywords: Narayana polynomial, colored tree, Stanley's bijective proof problem, Harer-Zagier formula.

(Concerned with sequences A001263, A006318, A035309 and A132393.)

Received September 13 2016; revised versions received January 18 2017; January 23 2017; January 25 2017. Published in *Journal of Integer Sequences*, January 26 2017.

Return to Journal of Integer Sequences home page.